

الباب الثاني

الإطار النظري

أ. فهم التقويم و تحليل بنود الأسئلة

1. فهم التقويم

التقويم هو عملية جَمْع البيانات أو الإعلام وتحليلها وترجمتها (له الكمية و الكيفية). يرتبط التقويم بالواقع و الموقف المستخدمين في صنع القرار.¹ يحصل التقويم على الإعلام لاختيار الحلّ الأحسن. يعرفُ التقويمُ درجةً نجاحِ المنهاج.² يُقوّمُ التقويمُ الأنشطة بالمرتجل والمخطّط والمنظّم والموجّه إلى الأهداف الواضحة.³ يُزْمَعُ التقويمُ العملية القيّمة على كلّ شيءٍ في التربية.⁴ يتضمّنُ التقويمُ بيانات العدد و غير العدد.⁵

¹Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 4

²Acep Hermawan, *Metodologi Pembelajaran Bahasa Arab*, Bandung: PT Remaja Rosdakarya, 2011, h. 276

³Chabib Thoha, *Teknik Evaluasi pendidikan*, Jakarta: Raja Grafindo Persada, 1994 h. 1

⁴Wayan Nurkancana dan Sumartana, *Evaluasi Pendidikan*, Surabaya: Usaha Nasional, cet. ke 4, h. 1

⁵M. Ainin dkk, *Evaluasi dalam Pembelajaran Bahasa Arab*, Malang: Misykat, cet. ke-1, 2006, h. 3

تقويم الشيء في اللغة أي إعطائه قيمة، وقوم الشيء وزنه، وفي التربية تقويم المعلم لأداء التلاميذ يعني إعطاء هذا الأداء قيمة ووزنا، وهو بصورة عامة يتضمن إصدار الحكم على الأشياء أو الأشخاص أو الأفكار.⁶

ومثل ذلك التعريف يعرف الديب التقويم على أنه عملية لإصدار أحكام والوصول إلى قرارات حول قيمة خبرة من الخبرات، وذلك من خلال التعرف على نواحي القوة والضعف فيها على ضوء الأهداف التربوية المقبولة بقصد تحسين عملية التعليم والتعلم.⁷

أما أهداف التقويم: الأهداف العامة و الخاصة

أ. الأهداف العامة

هدفُ التقويم عامًا هو يعرف المؤثرَ و فعاليةَ نظامِ التعلمِ عَنْ أهدافِ المادّةِ والطريقةِ والوسائلِ ومصدّرِ الدُّرُوسِ والبيئةِ أو نظامِ القيمةِ.⁸ يهدف

⁶ فاروق عبد السلام وآخرون، مدخل إلى القياس التربوي والنفسي، بريطانيا: دار إيلاف، 1994، ص. 16

⁷ أوريل بحر الدين، مهارات التدريس نحو إعداد مدرّس اللغة العربية الكفاء، مالانق: إن- مالك برس، 2011، ص. 197

⁸ Zainal Arifin, *Evaluasi Pembelajaran*, Bandung: PT Remaja Rosdakarya, 2011, h. 15

التقويمُ لِأَسَاسِ تَغْيِيرِ الْمُنْهَاجِ أَيْضًا. وَ يَكُونُ تَغْيِيرُ الْمُنْهَاجِ سَدِيدًا حِينَما تَغْيَرُ ذَلِكَ الْمُنْهَاجُ عَلَى أَسَاسِ نَتَائِجِ التَّقْوِيمِ بِالْمَقْدَارِ الْأَكْبَرِ.⁹

ب. الأهداف الخاصة

أهداف التقويم خصوصاً كما يلي :

1. فِي التَّعْلِيمِيَّةِ، يُثَبِّتُ التَّقْوِيمُ كَفَاءَةَ التَّعْلُمِ الْمَعْيَنَةِ عَلَى كُلِّ الطَّلَّابِ وَ يُصَلِّحُ التَّقْوِيمُ عَمَلِيَّةَ التَّعْلُمِ وَ التَّعْلِيمِ.
2. فِي نَتِيجَةِ التَّعْلُمِ، يَعْرِفُ التَّقْوِيمُ فَرْقَ قُدْرَةِ الطَّلَّابِ فِي اسْتِيعَابِهِمْ وَ يَقَيِّسُ التَّقْوِيمُ نَجَاحَ الطَّلَّابِ بِالشَّخْصِيَّةِ أَوْ الْجَمْعِيَّةِ.
3. فِي التَّشْخِيصِ وَ سَعْيِ التَّحْسِينِ، يَعْمَلُ التَّقْوِيمُ التَّشْخِيصَ إِلَى مَصَاعِبِ تَعْلُمِ الطَّلَّابِ، ثُمَّ يَحْصِلُ التَّقْوِيمُ عَلَى التَّحْسِينِ إِلَى عَمَلِيَّةِ التَّعْلُمِ.
4. فِي الْوَضْعِيَّةِ، يَخْصُلُ التَّقْوِيمُ عَلَى الْإِعْلَامِ عَنْ قُدْرَةِ الطَّلَّابِ حَتَّى وَضْعِهِمْ بِنَاءً عَلَى مَوْهَبَتِهِمْ وَ رَغْبَتِهِمْ.
5. فِي الْمُنْتَحِيَّةِ، يُسْتَخْدَمُ التَّقْوِيمُ كَالَّةِ الْمُنْتَخَبِ فِي قَبُولِ الطَّلَّابِ، ثُمَّ يُسْتَخْدَمُ حَاصِلُ التَّقْوِيمِ لِلتَّوْفِيقِ وَ التَّوْعِيَّةِ

⁹Sukardi, *Evaluasi Pendidikan Prinsip dan Operasionalnya*, Jakarta: Bumi Aksara, 2011, h. 9-10

6. في التوفيقيّة والاستشاريّة، يُساعد التقيّم الطلاب لاختيار التربيّة الآتية،

و لاختيار أنشطة مناهج ولّوئيّّة دراسيّة وغيرها.

7. في المنهج، يقيس تقويم المنهج على نجاح المؤسّسة التربويّة حتّى يمكن

التقويم أن يعطي القيمة إلى مؤثّر المؤسّسة التربويّة.¹⁰

ويجب أن تركز خطة تطوير التقويم التربوي على مجموعة من الأهداف

أهمها:

1. تطوير البنية الأساسيّة للأجهزة والإدارات المختصة بالتقويم ودعمها

بالكفاءات الفنيّة المؤهلة.

2. تطوير النظم والإجراءات الفنيّة والعمليّة المتبعة في إعداد الاختبارات

وبنائها وحفظها وتصحيحها والاستفادة من نتائجها.

3. تحسين نوعية الاختبارات المدرسية من حيث مستوى إعدادها وتطبيقها

وأوجه الاستفادة منها.

4. مراجعة نتائج الاختبارات والتقويم وتعديلها في ضوء الأسس القياسية

وبما يتفق مع الغرض منها.

¹⁰Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitab, 2012, h. 9-19

5. سد الاحتياج من الاختبارات المدرسية اللازمة لعمليات التقويم التربوي.

6. سد الاحتياج اللازم من أدوات التقويم في مجال اختبارات الاستعداد

والميل والاتجاهات وغيرها من الاختبارات المناسبة للتشخيص

والإرشاد.¹¹

أما وظائف التقويم في التربية والتعليم يعني: أولاً، يعرف التقويم تقدّم

الطلاب وتطوّرهم ونتيجتهم بعد عملية التعلّم مادام الوقت. ثمّ تُستخدم

نتائج التقويم للتّحسين في طريقة تعلّم الطلاب (الوظيفة البادئة). ثانياً،

يعرف التقويم درجة نتيجة التدريس.¹²

2. فهم تحليل بنود الأسئلة

تحليل بنود الأسئلة هو الاصطلاح المستخدم لازمّاع الحساب و قياس إجابة

المرؤوس إلى البنود. عامّة، يهدف تحليل البنود أن يُزْمَعَ هل البنود الجيدة أو البنود

القيّحة كآلة القياس، حتّى يمكن تحليل البنود أن يُقَصَّر أو يُطَوَّل الاختبار و يزيد

¹¹ رائد خليل العبادي، الاختبارات المدرسية، عمان: مكتبة المجتمع العربي للنشر والتوزيع، 2005، ص.

¹²Ngalm Purwanto, *Prinsip-prinsip dan Teknik Pengajaran*, Bandung: PT. Remaja Rosdakarya, 2000, h.5

الصدق أو ثباته. يعرف تحليل الأسئلة وظيفّة الأسئلة. يُعملُ التحليل بالطريقتين،

يعني: التحليل الكميّ و التحليل الكيفي.¹³

ب. فهم الاختبار

قيل "الاختبار" يُأخذُ من اللغة الفرنسيّة "*Testum*" يعني الصحن الذي

يُستخدم ليُفرد المعادن الخاصة (الذهب والفضّة). يعرف جاميس م س. كاتيل هذا

الاصطلاح إلى المجتمع في السنة 1980 في كتابه "*Mental Test and Measurement*".

ويزدهم هذا الاصطلاح في الولايات المتحدة الأمريكية.¹⁴

الاختبار هو طريقة أو آلة للحصول على تقويم السؤال. يحصل السؤال على

البيانات (القيمة) عن سلوك أو منجز الطلاب. ثم تُقابل تلك القيمة بقيمة الطلاب

الآخر أو بقيمة المعيار. لغة، اختبار اللغة هو سؤال الذي يُجيبه الطلاب بالأهداف

لقياس درجة استيعاب مهارة اللغة، ويوضح اختبار اللغة نشأة المنجز ثم تُقابل تلك

نشأة المنجز بمنجز الطلاب الآخر.¹⁵ اختبار نتيجة التعلم هو الأسئلة و الوظائف التي

¹³Sumarna Supranata, *Analisis, Validitas, Reabilitas, dan Interpretasi Hasil Tes*, Bandung: PT Remaja Rosdakarya, 2004, h. 1

¹⁴Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: PT Bumi Aksara, 2009, h. 10

¹⁵Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: PT Bumi Aksara, 2009, h. 168

تجب على كلّ الطلاب إجابتها لقياس تقدّم تعلّم الطلاب. و يحصل هذا الاختبار على البيانات الكميّة.¹⁶

أمّا الوظائف من الاختبار هي: أولاً، يقيس الاختبار على الطلاب (آلة القياس). و يقيس الاختبار درجة تطوّر أو تقدّم الطلاب بعد عمليّة التعلّم و التعليم مادام الوقت. ثانياً، يقيس الاختبار نجاح التدريس، لأنّ الاختبار يعرف استيعاب نجاح التدريس.¹⁷

ج. أشكال الاختبار

1. الاختبار الشخصي (غير الموضوعي). يُعطى الاختبار للطلاب في العدد القليل بين 10-50 سؤالاً. تطلب الأسئلة المقالة إلى الطلاب للمنظمة والتفسير وإرتباط الفهم. اختبار الأسئلة المقالة له فضيلة وهي: يُسهّل الاختبار في العِدّة والترتّب، ولا يُعطى فرصة للطلاب في المخاطر. يمكن اختبار الأسئلة المقالة أن يقيس كفاءة الطلاب في فهم المشكلة، ولكن اختبار أسئلة المقالة له درجة الصدق المنخفض وثباته والفحص الشخصي.

¹⁶Slameto, *Evaluasi Pendidikan*, Jakarta: PT Bumi Aksara, 1999, h.30

¹⁷Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada, 2007, h. 67

2. الاختبار الموضوعي أو اختبار الإجابة القصيرة، يُفحص اختبار الموضوع بالموضوعية و يُعطى في العدد الكثير. الاختبار الموضوعي له فضيلة وهي: يُمثّل الاختبار الموضوعي المادة الواسعة، والموضوعية، والأسهل في الفحص، لأنه يمكن أن يتنفع التكنولوجيا و شخص آخر في فحصه. الاختبار الموضوعي له نقيصة أيضا وهي: يُرتّب الاختبار الموضوعي صعوبة من اختبار الأسئلة المقالة، وتحمين (لديه فرصة بالنصيب) ويفتح الاختبار الموضوعي فرصة إلى الطلاب للتعاون. يتضمن الاختبار الموضوعي إلى الاختبار الخطأ والصواب، والاختبار من متعدد، واختبار المزاوجة و الاختبار التكميلي.¹⁸

أما جنس الاختبار هو:

1. اختبار الوضع. تكون نتيجة اختبار الوضع أساسا في مُعطِ الدّرس والتوفيق أو بنوءِ قدرة الطلاب على استيعاب المادة.
2. اختبار التشخيص. يُستخدم اختبار التشخيص لتعريف صعوبة تعلّم الطلاب. يُعمل اختبار التشخيص بناءً على نتيجة اختبار البادئة.
3. اختبار البادئة. يُعمل اختبار البادئة في عمليّة التعلّم، وخصوصا في آخره.

¹⁸Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: PT Bumi Aksara, 2009, h. 168

4. اختبار التلخيص. يُعمل اختبار التلخيص في اختبار امتحان الآخر. يهدف اختبار

التلخيص لتعريف مدى استيعاب الطلاب على أهداف التعلم. تمكن نتيجة اختبار

التلخيص أن تكون أساساً في تحسين المنهاج في الوقت الطويل.¹⁹

د. الاختبار الموضوعي في الاختبار من متعدد

يتكوّن الاختبار من متعدد على البيان أو المعرفة عن شيء مُبهم. يكمل

الاختبار من متعدد باختيار الإجابة. يتكوّن المتعدد على جزء البيان واختيار الإجابة.

أما اختيار الإجابة يتكوّن على الإجابة الصحيحة والانتباه (*distractor*).²⁰ تُستخدم

بنود الاختبار من متعدد أن يقيس المعيار أو فهم العلم الواضح.²¹

وأما خصائص الاختبار من متعدد الجيد كما يلي:

1. باستخدام سؤال أو بيان

بيان	سؤال
* ليس معي قلم مرسمة أ. و	* ما معنى البائع؟ أ. Petani

¹⁹ Nana Sudjana, *Dasar-dasar Penilaian Hasil Belajar*, Bandung: 1986, h.25-29

²⁰ Nana Sudjana, *Dasar-dasar Penilaian Hasil Belajar*, Bandung: 1986, h.25-29

²¹ Sukardi, *Evaluasi pendidikan Prinsip dan Operasionalnya*, Jakarta Timur: Bumi Aksara, 2011, cet. ke 5, h.125

ب. أو	ب. Pedagang
ج. ثمّ	ج. Polisi
د. بل	د. Tukang Sapu

2. يجب أن تكون الأجوبة البديلة متسلسلة بمنطقيّة و منهجية

الأجوبة الجيدة	الأجوبة غير الجيدة
* الطلاب في دروسهم	* تذهب عائشة إلى المستشفى في الساعة
أ. يجتهد	أ. يجتهد
ب. يجتهدان	ب. يجتهدون
ج. يجتهدون	ج. أجتهد
د. تجتهد	د. نجتهد

3. الأجوبة توجه إلى المقصود و إنما يجوز أن تطول الأسئلة

الأجوبة الجيدة	الأجوبة غير جيّدة
* أذهب إلى المدرسة صباحا أرجع منها بعد الظهر	* أذهب إلى المدرسة صباحا

أ. أو أرجع منها بعد الظهر	أ. أو
ب. بل أرجع منها بعد الظهر	ب. بل
ج. و أرجع منها بعد الظهر	ج. و
د. ثم أرجع منها بعد الظهر	د. ثم

4. المعنى الرئيسى للسؤال لا ينفصل بالفراغ

السؤال الجيد	السؤال غير جيد
* الشرطي المرور	* أي الشرطي المرور
أ. يفحص	أ. يفحص
ب. ينقل	ب. ينقل
ج. يصنع	ج. يصنع
د. ينظم	د. ينظم

5. الأجوبة متعددة الخيارات متناسقة في تركيبها (الجملة/العبارة/الكلمة).

الأجوبة الجيدة	الأجوبة غير جيدة
----------------	------------------

* ما معنى "الزِّي المدرسي" ؟	* ما معنى "الزِّي المدرسي" ؟
أ. Memakai	أ. Perlengkapan Sekolah
ب. Menggunakan	ب. Jadwal Sekolah
ج. Seragam Sekolah	ج. Lambang Sekolah
د. Kewajiban Sekolah	د. Seragam Sekolah

6. الأجوبة البديلة المتجانسة فيها تحويلات، محولات، خدعات

الأجوبة غير جيّدة	الأجوبة الجيّدة
* من هو الذي يزرع النباتات ؟	* من هو الذي يزرع النباتات ؟
أ. الفلاح	أ. الفلاح
ب. المهندس	ب. المهندس
ج. البائعون	ج. البائع
د. الطبيب	د. الطبيب

7. لا تحتفظ "الشّفرة" للحصول على الأجوبة الصحيحة

السؤال غير جيّد	السؤال الجيّد
<div>Gambar tukang sapu menyapu jalanan</div> <div>* ما هذه الصورة ؟</div>	<div>Gambar tukang sapu menyapu jalanan</div> <div>* ما هذه الصورة ؟</div>

<p>أ. الفلاح يزرع النباتات</p> <p>ب. الكناس يكنس الشارع</p> <p>ج. الطبيب يفحص المريض</p> <p>د. السيارة تنقل البضائع</p> <p>ثم في السؤال الآت : * الكناس ... شوارع</p> <p>أ. يحصل</p> <p>ب. يجمع</p> <p>ج. ينقل</p> <p>د. يكنس</p> <p>(هذا السؤال، في السؤال السابق تُحْتَفَظُ "الشِّفْرَة" للحصول على الأجوبة الصحيحة في السؤال الآت)</p>	<p>أ. الفلاح يزرع النباتات</p> <p>ب. الكناس يكنس الشارع</p> <p>ج. الطبيب يفحص المريض</p> <p>د. السيارة تنقل البضائع</p> <p>ثم في السؤال الآت : * ما معنى الفلاح يزرع النباتات ؟</p> <p>أ. Dokter memeriksa Pasien</p> <p>ب. Petani menanam sayur</p> <p>ج. Tukang sapu menyapu jalanan</p> <p>د. Pemulung memungut sampah</p> <p>(ينفصل بين البنود والبنود الأخرى إمّا البنود السابقة أو البنود الآتية، حتّى لا تُحْتَفَظُ "الشِّفْرَة" للحصول على الأجوبة الصحيحة إمّا في السؤال أو الأجوبة البديلة)</p>
---	---

8. تفترق كثير طوال الأجوبة بين واحدة وغيرها

الأجوبة الجيدة	الأجوبة غير جيدة
* يجب أحمد	* يجب أحمد
أ. كتابة	أ. كتابة وقراءة
ب. قراءة	ب. قراءة
ج. سباحة	ج. سباحة و رياضة
د. رياضة	د. رياضة

9. لا تستخدم فيها كلمة "كل الأجوبة صواب" أو "كل الأجوبة خطأ"

الأجوبة الجيدة	الأجوبة غير جيدة
* الطالب يذاكر	* الطالب يذاكر
أ. دروسه	أ. دروسه
ب. دروسهما	ب. دروسهما
ج. دروسهم	ج. دروسهم
د. دروسها	د. الاجوبة السابقة كلّها صحيحة

أو.....	
د. الاجوبة السابقة كلها خطأ	

10. لكل سؤال هناك جواب الصحيح الواحد

الأجوبة الجيدة	الأجوبة غير جيدة
* ما عاصمة من الإندونيسيا ؟	* ما عاصمة من الإندونيسيا ؟
أ. جاكرتا	أ. جاكرتا
ب. سورابايا	ب. سورابايا
ج. بالنكا رايا	ج. بالنكا رايا
د. مالوكو	د. جاكرتا

11. متسقة في جعل شكل الاختيار.²²

الأجوبة الجيدة	الأجوبة غير جيدة
١.....	١.....
أ.	أ.
ب.	ب.

²² Nurul Wahdah, *Bentuk-bentuk Tes Bahasa Arab*, Bahan ini disampaikan dalam Workshop Bahasa Arab, 26 November 2011, h.4

ج. ٠	ج. ٠
د. ٠	د. ٠
٢ ٠	2 ٠
أ. ٠	أ. ٠
ب. ٠	ب. ٠
ج. ٠	ج. ٠
د. ٠	د. ٠

12. لا يحتفظ "الخطأ" في كتابة السؤال أو الأجوبة البديلة الصحيحة

السؤال غير جيّد	السؤال الجيّد
* الشرطي المُرور	* الشرطي المُرور
أ. يَفْحَصُ	أ. يَفْحَصُ
ب. يَصْنَعُ	ب. يَصْنَعُ
ج. يُنْظَمُ	ج. يُنْظَمُ
د. يَبِيعُ	د. يَبِيعُ

هـ. معيار الاختبار الجيد في اختبار اللغة

الاختبار الجيد له صدق وثبات الاختبار، و الدرجة صعوبة، و قدرته تفرقه.²³

1. الصدق

يُعرف الصدق بالسديد والصحيح والمضبوط.²⁴ يحيلُ صدقُ الاختبار

(الصحة *validity*) إلى قدرة الاختبار في قياس خصائص القاس.²⁵ يكون اختبار

نتيجة التعلم صدقًا إذا كان له سديد في قياس نتائج تعلم الطلاب. تُفرّق قدرته

ودرجة الصدق بالمعيار "منخفض جدًا ومنخفض ومقبول وجيد وجيد جدًا".

يُدلُّ الصدق على النسبة المئوية (مثل 65%) أو فهرس العلاقة (مثل 0,65).²⁶

يُفرّق الصدق في الأجناس وهي:

أ. صدق المحتوى (*content validity*)

ب. صدق التركيب (*construct validity*)

ج. صدق النبوءة (*predictive validity*)

²³ Shodiq Abdullah, *Evaluasi Pembelajaran*, Semarang: Pustaka Rizki Putra, 2012, h.98

²⁴ Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 176

²⁵ Acep Hermawan, *Metodologi Pembelajaran Bahasa Arab*, Bandung: PT Remaja Rosdakarya, 2011, h. 277

²⁶ Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 176-177

د. صدق المترامنة (concurrent validity).²⁷

2. الثبات

قيل الثبات بمضبوط الاختبار الذي عزا إلى قدرة الاختبار، وتستطيع تلك قدرة الاختبار أن تقيس الصفات بالمتوافق. وعلى تركيب الاختبار الجيد له تلك الشرط حتى يستطيع الاختبار أن يوضح الحالة الحقيقة إلى المشترك الاختبار.²⁸

ويقال الاختبار له صفة الثبات إذا كان له المتوازن في القياس إلى موضوع النفس. عندئذ، فإذا الاختبار له قيمة أو بيانات نتائج امتحان الطلاب متوازنا فكان الاختبار ثابتا، أو متى وأين ومن الذى يعمل أو يفحص أو يقوم ذلك الاختبار متوازنا فكان الاختبار ثابتا أيضا.²⁹

3. الموضوعية

ويقال الاختبار موضوعيا، إذا كان له تركيب وعمل بحالة حقيقة في مادة الاختبار. إنَّ الحلة الحقيقة هي المادة التي تُعطى وفقًا للمؤشر من المدرس إلى

²⁷Sumarna Surapranata, *Analisis, Validitas, Reliabilitas, dan INTERPRETASI HASIL TES Implementasi Kurikulum 2004*, bandung: PT Remaja Rosdakarya, cet 4, 2009. Hal 50-51

²⁸Acep Hermawan, *Metodologi Pembelajaran Bahasa Arab*, Bandung: PT Remaja Rosdakarya, 2011, h. 277

²⁹Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabiah, 2012, h. 178

الطلاب بناءً على معيار المؤشر. يشدد الثابت سديداً في نتيجة الاختبار ولكن

تشدد الموضوعية سديداً في موضوع الاختبار و القيمة.³⁰

4. العملية

الاختبار العملي هو اختبار سهل في عمليته، ولا يطلب الاختبار العملي

المرافق المعقد. ينبغي على المدرس أن يختار الاختبار السهل في تقييمه. وكذلك

سهولة التقييم فينبغي عليه أن يتم بالمبادئ التوجيهية للتقويم مثل المقياس لكل

من جنس الاختبار.³¹

5. التفريقية

لا يجوز أن يكون الاختبار سهلاً جميعاً أو صعباً جميعاً. الاختبار الجيد له

درجة صعوبة و قدرة التفريق بين التلميذ الماهر و التلميذ الجاهل. ثمّ لتعريف

مدى الاختبار له قدرة التفريق بين الفرقتين فأعمل تحليل قدرة تفريق الاختبار

باستخدام الرموز المؤكدة.³²

³⁰Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 178-179

³¹Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 179-180

³²Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 180

و. تحليل جودة بنود اختبار اللغة

1. تقييم الاختبار المصنوع الذاتي :

لتقييم جودة الاختبار من خلال: يحلل الباحث عن الأسئلة بالصدق، ويراجع

الباحث بنود الأسئلة، ويراجع الباحث الصدق والثبات.³³

2. تحليل بنود الأسئلة

يهدف تحليل الأسئلة أن يعرف الأسئلة الجيدة و القبيحة. يحصل تحليل

الأسئلة على المعلومات عن الأسئلة غير جيدة وتحسين الأسئلة.

أ. تحليل صدق الاختبار

يرتبط الصدق بالاختبار. يُفَرَّق صدق الاختبار بالجنسين وهما: صدق

المنطق وصدق التجريب. يُشَبَّه صدق المنطق بالتحليل الكيفي إلى الأسئلة. تُزَمَّع

وظيفة الأسئلة من التحليل الكيفي بناء على محكّ المادّة والتركيب واللغة.³⁴

³³Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 181-182

³⁴Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 1999, h. 61

(1) إثباتُ صدقِ المحتوى

يعمل صدق المحتوى بتركيب محكّ الاختبار، ثم يقابل محكّ الاختبار بمحتوى بنود الأسئلة. يحلّل محكّ الاختبار لتعريف أسئلة اختبار الصدق مدى أسئلة الاختبار هل تمثّل جميع المادّة. ولذلك، لا يملك صدقُ محتوى الاختبار الوزنَ في حساب الإحصاءات، ولكن قيل أنّ محتوى الأسئلة صدقًا حينما تملك الأسئلة خصائص الاختبار من متعدّد الجيّد و إنسجامًا بالمؤشّر والمادّة واستخدام لغة الأسئلة.³⁵

(2) إثباتُ صدقِ التركيب

يُستخدم صدق التركيب لأدوات مقياس متغيّر التركيب. صدق التركيب له العمليّة الخاصّة مثله: كأدوات مقياس الموقف والرغبة ومفهوم النفس وأسلوب القيادة وتعليل المنجز وغيرها، وله العمليّة المنتهيّة أيضًا مثله: كأدوات مقياس الموهبة (اختبار الموهبة) والذكاء وذكاء العاطفيّة وغيرها. يَجْبُرُ صدقُ التركيبُ البُعْدَ والمؤشّرَ بالرموز أو بالإعتبار إلى:

³⁵ Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 189-190

أ) هل المؤشّر و التركيب سديد

ب) على المؤشّر من التركيب أن يكون متجانسا ومتوافقا ومتقاربا كي

يستطيع المؤشّر أن يقيس أدوات القاس.

ج) على المؤشّر أن يكون كاملا كي يستطيع المؤشّر أن يقيس التركيب

بالتامة.³⁶

3) إثباتُ صدق البنوءة

صدق البنوءة هو نفس صدق التجرب، ويُزَمَعُ صدقُ البنوءة بالبنوءة

الداخلية والخارجية. أمّا البنوءة الداخلية هي اختبار أو أدوات بنفسها،

والبنوءة الخارجية هي نتائج قياس الأدوات أو الاختبار الآخر من خارج

الأدوات.

أ) الصدق الداخلي

يُقاسُ الصدقُ الداخلي بالوزن الذي يستخدم الأدوات

كمّحدة (جميع البنود) لازماع صدق البنود أو لازماع بنود من

الأدوات. إذا كان معامل العلاقة بين قيمة البنود و قيمة جميع الأدوات

³⁶ Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 191

مُوجِباً أو مهمّماً، فيكون البنودُ صدقاً بناءً على مكيال الصدق الداخليّ. وإذا كان وزنُ معاملِ العلاقة بين قيمة البنود و قيمة الجميع قيمةً موجبةً وكلّما يَكْبُرُ معاملُ علاقتهما، فيكون صدق البنود مرتفعاً. يوضّح معاملُ العلاقة المرتفعة بين قيمة البنود و قيمة الجميع على ارتفاع المتوفّق بين نتائج قياس جميع الأدوات و نتائج قياس بنود الأدوات. أمّا استخدام الرمز في حسابه من خلال (*product moment*) بعددٍ خَشِنٍ.³⁷

ب) الصدق الخارجيّ

كلّما تحصل الأدوات على معامل العلاقة المرتفعة فيكون صدق الأدوات الانتشار جيّده أيضاً. تُستخدم البنوءُ قيمة الفهرس r ($r_{\text{-tabel}}$) لامتحان الصدق الخارجيّ. إذا يكون معاملُ العلاقة بين قيمة نتائج قياس الأدوات الانتشار و قيمة نتائج قياس الأدوات الفصيح أكبر من $r_{\text{-tabel}}$ فتكون أدوات الانتشار صدقاً بناءً على بنوء خارجيّة الاختيار (نتائج قياس الأدوات الفصيح). إذن، يصمّم امتحان الصدق

³⁷ Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabiah, 2012, h. 191-194

عن الصدق أو غير صدق الأدوات كمتحدة، ليس صدق أو غير الصدق بنود الأدوات كمثل في الصدق الداخلي. يحسب الصدق الخارجي بصنع فهرس الاستعداد ثم يدخل رموز علاقة *product* *moment* في الصدق الخارجي.³⁸

ب. تحليل ثبات الاختبار

يُعمل تحليل الثبات بتام و يُكاشف تحليل الثبات بعدد معامل العلاقة بين +1,00 و -1,00. يفرق الثبات بنوعين، وهما:

1) ثبات متوافق الإصغاء

يُستخدم ثبات متوافق إصغاء في الاختبار أو الادوات لقياس موضوع القياس، ثم يُقاس الموضوع بنفس الآلة/الطريقة مرة أخرى، إذا حصل قياس الثاني غير الموافق، فلا توضح نتيجة القياس حالة الموضوع الحقيقة. يفحص ثبات إصغاء المجيب إلى الاختبار بثلاث الطرق:

³⁸ Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitab, 2012, h. 194-195

أ) طريقة المناظرة أو *equivalent method*

يُقاس ثبات الاختبار باختبارين معتدلين في الأهداف و درجة الصعوبة وتركيبهما، لكنهما مختلفان في بنود الأسئلة. ثم يُعطى الاختباران المعتدلان إلى المجيب أو موضوع الاختبار في وقت واحد. تُعلّق القيمة من فرقتين بنود الاختبار لمعرفة ثبات الاختبار. بهذه طريقة المناظرة، يُعلّق الاختباران المتوازن (مثل اختبار اللغة العربية في مجلّد "أ" و مجلّد "ب") إلى فرقة الطلاب بنفس لمعرفة الثبات. يدلّ معامل العلاقة من نتيجتي الاختبار معامل ثبات الاختبار "أ".

ب) طريقة الاختبار-إعادة الاختبار أو *test-retest method*

يملك المختبر (صانع الاختبار) مجلّد الاختبار فقط، ولكن ذلك الاختبار يُمتحن مرّتين. ثمّ يحسب المختبر النتائج من مرّتين ذلك الاختبار في العلاقة. ويكون اختبار الثاني عاما أحسنا من اختبار الأولى. بل هنا الحال غير المشكّلة، لأنّ المهمّ قد يحصل من الاختبارين تناسق العلاقة صدقها.

ج) طريقة تقسيم النصف *split-half method*

تُقاس طريقة تقسيم النصف بفرقتين البنود المتوازن في نفس الوقت. تكون فرقة البنود نصفًا من جميع الاختبار فتأخذ فرقة البنود الأولى من بنود الاختبار بالرقم الوتر وفرقة البنود الثانية من بنود الاختبار بالرقم الشفع. لتعريف الثبات من جميع الاختبار فتستخدم

طريقة تقسيم النصف برموز Spearman-Brown.³⁹

2) ثبات متوافق انضمام البنود

يرتبط ثبات متوافق انضمام البنود بالمتوافق بين بنود الاختبار. إذا تكون نتيجة القياس بين احد البنود والأخرى في موضوع القياس نفس غير المتوافق، فليس موضوع القياس خطأ ولكن آلة القياس (الاختبار) خطأ. يقال ليس للاختبار صفة الثبات إلى موضوع القياس أو قيل أنّ الاختبار ثباته منخفضاً.⁴⁰ يُحسبُ ثبات الأدوات برموز Kuder Richardson و

³⁹Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 197-198

⁴⁰Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 201

Anova و KR-20 KR-21 و Spearman Brown و Alpha Cronbach

⁴¹.Hayt

ج) تحليل درجة الصعوبة

درجة الصعوبة للأسئلة هي فرصة الإجابة الصحيحة على الأسئلة في

درجة القدرة. وتتكوّن درجة الصعوبة على الفهرسة الصعوبة (*difficulty*)

(*index*).⁴² أمّا الفهرسة الصعوبة تدلّ على نسبة الطلاب بالإجابة

الصحيحة.⁴³ كلما تكبر درجة الصعوبة ونتيجة الحساب فتكون الأسئلة

سهولة.⁴⁴

تحلّل درجة الصعوبة إلى بنود الاختبار لتعريف الصعوبة أو سهولة

الاختبار، إمّا الاختبار الإجمالي (*overall test*) أو كلّ بنود الاختبار. تُحسبُ

درجة الصعوبة من المقارنة بين عدد من الطلاب بإجابة الأسئلة الصحيحة

وبإجابة الأسئلة الخطأ. ولذلك، تُحسبُ درجة صعوبة بنود الاختبار بتقسيم

عدد من الطلاب بإجابة الأسئلة الصحيحة على عدد مشترك الاختبار.

⁴¹ Ridwan, *Metode dan Teknik Menyusun Tesis*, Bandung: Alfabeta, 2010, h. 113

⁴² Kusaeri Suprananto, *Pengukuran dan Penilaian Pendidikan*, Yogyakarta: Graha Ilmu, 2012, h.174

⁴³ Suke Silverius, *Evaluasi Hasil Belajar dan Umpan Balik*, Jakarta: PT Grasindo, 1991 h. 166

⁴⁴ Suke Silverius, *Evaluasi Hasil Belajar dan Umpan Balik*, Jakarta: PT Grasindo, 1991 h. 166

وتكون البنود جيّدة حينما درجة الصعوبة لها معتدلة هي غير صعبة جدًا و غير سهلة جدًا. ولهذا السبب تملك درجة صعوبة البنود العلاقة بقدرة التفريق. إذا كانت البنود لها درجة الصعوبة الأقصى، فتكون قدره التفريق منخفضا. وإذا كانت البنود سهلة جدًا، فليس لها قدرة التفريق. لذلك، تتمسك درجة الصعوبة الأسئلة حدّ قدرة التفريق.⁴⁵

(د) تحليل قدرة التفريق

يقيس حساب قدرة التفريق بنود الأسئلة لتفريق الطلاب، هل الطلاب قد استوعبوا أو لم يستوعبوا الكفاءة بناءً على البنوءة. كلما يرتفع معامل قدرة التفريق في بنود الأسئلة، فتكون بنود الأسئلة استطاعا للتفريق بين طلاب الذين استوعبوا وغيرهم.⁴⁶

فهرس التفريق هو العدد الذي يدلّ على كبير قدرة التفريق.⁴⁷ يتّزحزح فهرس التفريق (قدرة التفريق) بين 0,00 حتى 1,00. يستخدم فهرس التفريق علامة سلبية، و تُستخدم العلامة السلبية لتعريف جودة مُشترك الامتحان

⁴⁵Daryanto, *Evaluasi Pendidikan*, Jakarta: PT Rineka Cipta, 2008, h. 180-182

⁴⁶Zainal Arifin, *Evaluasi Pembelajaran*, Bandung: PT Remaja Rosdakarya, 2012, h. 273

⁴⁷Suharsimi Arikunto, *Dasar Dasar Evaluasi Pendidikan*, Jakarta: PT Bumi Aksara, 2009, h. 211

المقلوبة. المقلوبة هي تُفْتَرَضُ ولدا ماهرا إلى غير ماهر وبالعكس ذلك. وثلاث

النُّقْط في قدرة التفريق، وهي:⁴⁸

$$-1,00 = \text{قدرة التفريق سلبية}$$

$$0,00 = \text{قدرة التفريق منخفضة}$$

$$1,00 = \text{قدرة التفريق موجبة}$$

ز. التركيب و القياس

يستخدم القياس الكمي القياس الرمزي والترتيبي والنسبة والمسافة.⁴⁹ قيلت أنّ

الأسئلة صدقًا بالتجريبي حينما تملك تلك الأسئلة العلاقة على 0,39. وقيل أنّ

محتوى الأسئلة صدقًا حينما تملك الأسئلة خصائص الاختبار من متعدد الجيد و

إنسجامًا بالمؤشر والمادة واستخدام لغة الأسئلة. ثم قيلت أنّ درجة صعوبة الأسئلة

المعتدلة حينما تملك الأسئلة درجة الصعوبة بين 0,30 حتى 0,70. وقيلت أنّ قدرة

التفريق المعتدل حينما تملك الأسئلة قدرة التفريق على الحد الأدنى 0,19.

⁴⁸Moh Matsna dan Erta Mahyudin, *Pengembangan Evaluasi dan Tes Bahasa Arab*, Tangerang Selatan: Alkitabah, 2012, h. 184

⁴⁹Burhan Bungin, *Metodelogi Penelitian Kuantitatif*, Jakarta: PT Prenada Media Grup, 2009, h. 93

السؤال المقبول هو في:⁵⁰

المحكّ	المعامل	القرار
الصدق	$> 0,39$	مقبول
درجة الصعوبة	$0,30 - 0,70$	مقبول
قدرة التفريق	$> 0,19$	مقبول

⁵⁰ Sumarna Supranata, *Analisis Validitas, Reliabilitas, dan Interpretasi Hasil Tes Implementasi Kurikulum 2004*, Bandung: PT Remaja Rosdakarya, 2009, h. 47